

CRF Errors Corrected by the STIC Sys as Branch

CRF Processing Date: 8/12/2002  
 Edited by: [Signature]  
 Verified by: [Signature] (STIC staff)

Serial Number: 10/070,099

ENTERED

- ☐ Changed a file from non-ASCII to ASCII
- ☐ Changed the margins in cases where the sequence text was "wrapped" down to the next line.
- ☐ Edited a format error in the Current Application Data section, specifically: \_\_\_\_\_
- ☐ Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other \_\_\_\_\_
- ☐ Added the mandatory heading and subheadings for "Current Application Data".
- ☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.
- ☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically: \_\_\_\_\_
- ☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were: \_\_\_\_\_
- ☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited: \_\_\_\_\_
- ☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.
- ☐ Inserted colons after headings/subheadings. Headings edited included: \_\_\_\_\_
- ☐ Deleted extra, invalid, headings used by an applicant, specifically: \_\_\_\_\_
- ☒ Deleted: ☒ non-ASCII "garbage" at the beginning/end of files; ☐ secretary initials/filename at end of file; ☐ page numbers throughout text; ☐ other invalid text, such as \_\_\_\_\_
- ☐ Inserted mandatory headings, specifically: \_\_\_\_\_
- ☐ Corrected an obvious error in the response, specifically: \_\_\_\_\_
- ☐ Edited identifiers where upper case is used but lower case is required, or vice versa.
- ☐ Corrected an error in the Number of Sequences field, specifically: \_\_\_\_\_
- ☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.
- ☐ Deleted **ending** stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: \_\_\_\_\_
- ☐ Other: \_\_\_\_\_

#8



PCT10

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/070,099

DATE: 08/12/2002

TIME: 17:41:50

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08122002\J070099.raw

```

3 <110> APPLICANT: Niederweis Dr., Michael
4   Bossmann Dr., Stefan
6 <120> TITLE OF INVENTION: Method for the Production of a Channel-forming Protein
8 <130> FILE REFERENCE: 401172
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/070,099
C--> 11 <141> CURRENT FILING DATE: 2002-07-08
13 <150> PRIOR APPLICATION NUMBER: DE 199 43 520.0
14 <151> PRIOR FILING DATE: 1999-09-11
16 <150> PRIOR APPLICATION NUMBER: DE 199 41 416.5
17 <151> PRIOR FILING DATE: 1999-08-31
19 <160> NUMBER OF SEQ ID NOS: 9
21 <170> SOFTWARE: PatentIn Ver. 2.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 636
25 <212> TYPE: DNA
26 <213> ORGANISM: Mycobacterium smegmatis
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (1)..(636)
31 <223> OTHER INFORMATION: mspA-Gene
33 <400> SEQUENCE: 1
34 atg aag gca atc agt cgg gtg ctg atc gcg atg gtt gca gcc atc gcg   48
35 Met Lys Ala Ile Ser Arg Val Leu Ile Ala Met Val Ala Ala Ile Ala
36   1             5             10             15
38 gcg ctt ttc acg agc aca ggc acc tct cac gca ggc ctg gac aac gag   96
39 Ala Leu Phe Thr Ser Thr Gly Thr Ser His Ala Gly Leu Asp Asn Glu
40             20             25             30
42 ctg agc ctc gtt gat ggc cag gac cgc acc ctc acc gtg cag cag tgg  144
43 Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu Thr Val Gln Gln Trp
44             35             40             45
46 gac acc ttc ctc aat ggt gtg ttc ccc ctg gac cgc aac cgt ctt acc  192
47 Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp Arg Asn Arg Leu Thr
48             50             55             60
50 cgt gag tgg ttc cac tcc ggt cgc gcc aag tac atc gtg gcc ggc ccc  240
51 Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr Ile Val Ala Gly Pro
52 65             70             75             80
54 ggt gcc gac gag ttc gag ggc acg ctg gaa ctc ggc tac cag atc ggc  288
55 Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu Gly Tyr Gln Ile Gly
56             85             90             95
58 ttc ccg tgg tgg ctg ggt gtg ggc atc aac ttc agc tac acc acc ccg  336
59 Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe Ser Tyr Thr Thr Pro
60             100            105            110
62 aac atc ctg atc gac gac ggt gac atc acc gct ccg ccg ttc ggc ctg  384

```

## RAW SEQUENCE LISTING

DATE: 08/12/2002

PATENT APPLICATION: US/10/070,099

TIME: 17:41:50

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08122002\J070099.raw

```

63 Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr Ala Pro Pro Phe Gly Leu
64      115      120      125
66 aac tcg gtc atc acc ccg aac ctg ttc ccc ggt gtg tcg atc tcg gca 432
67 Asn Ser Val Ile Thr Pro Asn Leu Phe Pro Gly Val Ser Ile Ser Ala
68      130      135      140
70 gat ctg ggc aac ggc ccc ggc atc cag gaa gtc gca acg ttc tcg gtc 480
71 Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val Ala Thr Phe Ser Val
72 145      150      155      160
74 gac gtc tcc ggc gcc gag ggt ggc gtg gcc gtg tcg aac gcc cac ggc 528
75 Asp Val Ser Gly Ala Glu Gly Gly Val Ala Val Ser Asn Ala His Gly
76      165      170      175
78 acc gtg acc ggt gcg gcc ggc ggt gtg ctg ctg cgt ccg ttc gcc cgc 576
79 Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu Arg Pro Phe Ala Arg
80      180      185      190
82 ctg atc gcc tcg acc ggt gac tcg gtc acc acc tac ggc gaa ccc tgg 624
83 Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr Tyr Gly Glu Pro Trp
84      195      200      205
86 aac atg aac tga 636
87 Asn Met Asn
88      210
91 <210> SEQ ID NO: 2
92 <211> LENGTH: 1423
93 <212> TYPE: DNA
94 <213> ORGANISM: Mycobacterium smegmatis
96 <220> FEATURE:
97 <221> NAME/KEY: -10_signal
98 <222> LOCATION: (323)..(328)
99 <223> OTHER INFORMATION: putative promoter
101 <220> FEATURE:
102 <221> NAME/KEY: CDS
103 <222> LOCATION: (499)..(1134)
104 <223> OTHER INFORMATION: mspA-Gene
106 <220> FEATURE:
107 <221> NAME/KEY: RBS
108 <222> LOCATION: (492)..(496)
109 <223> OTHER INFORMATION: putative ribosome binding site
111 <400> SEQUENCE: 2
112 gttaacggag tcgggcccgtc gatacggcgg cgaagatcat ccggcagatt ggcgccctggt 60
114 taaacccgcg taaacactgg taccgcccgt ccgcgccgga aaaggttttg cctcacgggtg 120
116 aatatgtgac ctgaattgca cttcacgggt aaaagcggag gtaaccgcag gttgccgcag 180
118 caccctcaca gcttgggcca aggtgacgtg cagcgcacgc ctgccgggtgc cggatggcgg 240
120 tcaccgcaaa gtgtcaggca ctgccgaaag gtcagtcagc aaacttcaact gcggctgttg 300
122 tgcgaaagtgc ggttgtggga cgtatccgtt gctgccgcgc gccctggcgt ttatgtttct 360
124 gctgccaaact gtgagcagag cattagagac agatgtgatc ctcttagatc tccgaagtct 420
126 ctgaacagggt gttgagccgg ttgcagacaa caaacagggt gggcctgagg ggccgcgcggc 480
128 gatacagtta gggagaac atg aag gca atc agt cgg gtg ctg atc gcg atg 531
129      Met Lys Ala Ile Ser Arg Val Leu Ile Ala Met
130      1      5      10
132 gtt gca gcc atc gcg gcg ctt ttc acg agc aca ggc acc tct cac gca 579

```

## RAW SEQUENCE LISTING

DATE: 08/12/2002

PATENT APPLICATION: US/10/070,099

TIME: 17:41:50

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08122002\J070099.raw

```

133 Val Ala Ala Ile Ala Ala Leu Phe Thr Ser Thr Gly Thr Ser His Ala
134          15          20          25
136 ggc ctg gac aac gag ctg agc ctc gtt gat ggc cag gac cgc acc ctc 627
137 Gly Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu
138          30          35          40
140 acc gtg cag cag tgg gac acc ttc ctc aat ggt gtg ttc ccc ctg gac 675
141 Thr Val Gln Gln Trp Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp
142          45          50          55
144 cgc aac cgt ctt acc cgt gag tgg ttc cac tcc ggt cgc gcc aag tac 723
145 Arg Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr
146 60          65          70          75
148 atc gtg gcc ggc ccc ggt gcc gac gag ttc gag ggc acg ctg gaa ctc 771
149 Ile Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu
150          80          85          90
152 ggc tac cag atc ggc ttc ccg tgg tgg ctg ggt gtg ggc atc aac ttc 819
153 Gly Tyr Gln Ile Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe
154          95          100          105
156 agc tac acc acc ccg aac atc ctg atc gac gac ggt gac atc acc gct 867
157 Ser Tyr Thr Thr Pro Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr Ala
158          110          115          120
160 ccg ccg ttc ggc ctg aac tgg gtc atc acc ccg aac ctg ttc ccc ggt 915
161 Pro Pro Phe Gly Leu Asn Ser Val Ile Thr Pro Asn Leu Phe Pro Gly
162          125          130          135
164 gtg tgg atc tgg gca gat ctg ggc aac ggc ccc ggc atc cag gaa gtc 963
165 Val Ser Ile Ser Ala Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val
166 140          145          150          155
168 gca acg ttc tgg gtc gac gtc tcc ggc gcc gag ggt ggc gtg gcc gtg 1011
169 Ala Thr Phe Ser Val Asp Val Ser Gly Ala Glu Gly Gly Val Ala Val
170          160          165          170
172 tgg aac gcc cac ggc acc gtg acc ggt gcg gcc ggc ggt gtg ctg ctg 1059
173 Ser Asn Ala His Gly Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu
174          175          180          185
176 cgt ccg ttc gcc cgc ctg atc gcc tgg acc ggt gac tgg gtc acc acc 1107
177 Arg Pro Phe Ala Arg Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr
178          190          195          200
180 tac ggc gaa ccc tgg aac atg aac tga ttcttgacc gccgttcggt 1154
181 Tyr Gly Glu Pro Trp Asn Met Asn
182          205          210
184 cgctgagacc gcttgagatc ggcgcgtccc gctcccggtg tcgtcagctc atcggtgaca 1214
186 cgtgaactga cactcttctt agccggagcg kacgcgccga tcttgtgttc tgagcagttc 1274
188 tcagtcctgc cgccgcaaca ccagcgtgta cggcgtaacg agcctgccca ccaccgcgcg 1334
190 ccagggaacgc ccagcctgg gcaccacctc agcggtcggc acgatgcgcg gatcggtcac 1394
192 ctgaacgtc tcaccgttca tcaccgcgc 1423
196 <210> SEQ ID NO: 3
197 <211> LENGTH: 211
198 <212> TYPE: PRT
199 <213> ORGANISM: Mycobacterium smegmatis
201 <220> FEATURE:
202 <221> NAME/KEY: SIGNAL

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/070,099

DATE: 08/12/2002

TIME: 17:41:50

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08122002\J070099.raw

203 <222> LOCATION: (1)..(27)  
 204 <223> OTHER INFORMATION: putative signal sequence of the MspA protein  
 206 <220> FEATURE:  
 207 <221> NAME/KEY: PEPTIDE  
 208 <222> LOCATION: (28)..(211)  
 209 <223> OTHER INFORMATION: mature MspA-protein  
 211 <400> SEQUENCE: 3  
 212 Met Lys Ala Ile Ser Arg Val Leu Ile Ala Met Val Ala Ala Ile Ala  
 213 1 5 10 15  
 215 Ala Leu Phe Thr Ser Thr Gly Thr Ser His Ala Gly Leu Asp Asn Glu  
 216 20 25 30  
 218 Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu Thr Val Gln Gln Trp  
 219 35 40 45  
 221 Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp Arg Asn Arg Leu Thr  
 222 50 55 60  
 224 Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr Ile Val Ala Gly Pro  
 225 65 70 75 80  
 227 Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu Gly Tyr Gln Ile Gly  
 228 85 90 95  
 230 Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe Ser Tyr Thr Thr Pro  
 231 100 105 110  
 233 Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr Ala Pro Pro Phe Gly Leu  
 234 115 120 125  
 236 Asn Ser Val Ile Thr Pro Asn Leu Phe Pro Gly Val Ser Ile Ser Ala  
 237 130 135 140  
 239 Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val Ala Thr Phe Ser Val  
 240 145 150 155 160  
 242 Asp Val Ser Gly Ala Glu Gly Gly Val Ala Val Ser Asn Ala His Gly  
 243 165 170 175  
 245 Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu Arg Pro Phe Ala Arg  
 246 180 185 190  
 248 Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr Tyr Gly Glu Pro Trp  
 249 195 200 205  
 251 Asn Met Asn  
 252 210  
 258 <210> SEQ ID NO: 4  
 259 <211> LENGTH: 558  
 260 <212> TYPE: DNA  
 261 <213> ORGANISM: Artificial sequence  
 263 <220> FEATURE:  
 264 <223> OTHER INFORMATION: description of the artificial sequence: synthetic  
 266 <220> FEATURE:  
 267 <221> NAME/KEY: CDS  
 268 <222> LOCATION: (1)..(558)  
 269 <223> OTHER INFORMATION: synmspA-Gene  
 271 <400> SEQUENCE: 4  
 272 atg ggc ctg gac aac gaa ctg tcc ctg gtt gac ggc cag gac cgt acc 48  
 273 Met Gly Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr  
 274 1 5 10 15

## RAW SEQUENCE LISTING

DATE: 08/12/2002

PATENT APPLICATION: US/10/070,099

TIME: 17:41:50

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08122002\J070099.raw

```

276 ctg acc gtt cag cag tgg gac acc ttc ctg aac ggt gtt ttc ccg ctg 96
277 Leu Thr Val Gln Gln Trp Asp Thr Phe Leu Asn Gly Val Phe Pro Leu
278          20          25          30
280 gac cgt aac cgt ctg acc cgt gaa tgg ttc cac tcc ggt cgt gcg aaa 144
281 Asp Arg Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Lys
282          35          40          45
284 tac atc gtt gcg ggt ccg ggt gcg gac gag ttc gaa ggt acc ctg gaa 192
285 Tyr Ile Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu
286          50          55          60
288 ctg ggt tac cag atc ggc ttc ccg tgg tcc ctg ggt gtt ggt atc aac 240
289 Leu Gly Tyr Gln Ile Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn
290 65          70          75          80
292 ttc tct tac acc acc ccg aac atc ctg atc gac gac ggt gac atc acc 288
293 Phe Ser Tyr Thr Thr Pro Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr
294          85          90          95
296 gct ccg ccg ttc ggt ctg aac tct gtt atc acc ccg aac ctg ttc ccg 336
297 Ala Pro Pro Phe Gly Leu Asn Ser Val Ile Thr Pro Asn Leu Phe Pro
298          100          105          110
300 ggt gtt tct atc tct gct gat ctg ggc aac ggt ccg ggt atc cag gaa 384
301 Gly Val Ser Ile Ser Ala Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu
302          115          120          125
304 gtt gct acc ttc tct gta gac gtc tct ggt gct gaa ggt ggt gtt gct 432
305 Val Ala Thr Phe Ser Val Asp Val Ser Gly Ala Glu Gly Gly Val Ala
306          130          135          140
308 gtt tct aac gct cac ggc acc gtt acc ggt gcg gct ggc ggt gtt ctg 480
309 Val Ser Asn Ala His Gly Thr Val Thr Gly Ala Ala Gly Gly Val Leu
310 145          150          155          160
312 ctg cgt ccg ttc gct cgt ctg atc gct tct acc ggt gac tct gtt acc 528
313 Leu Arg Pro Phe Ala Arg Leu Ile Ala Ser Thr Gly Asp Ser Val Thr
314          165          170          175
316 acc tac ggt gaa ccg tgg aac atg aac tga 558
317 Thr Tyr Gly Glu Pro Trp Asn Met Asn
318          180          185
322 <210> SEQ ID NO: 5
323 <211> LENGTH: 185
324 <212> TYPE: PRT
325 <213> ORGANISM: artificial sequence
327 <220> FEATURE:
328 <221> NAME/KEY: PEPTIDE
329 <222> LOCATION: (1)..(184)
330 <223> OTHER INFORMATION: rMspA
332 <220> FEATURE:
333 <223> OTHER INFORMATION: description of the artificial sequence: synthetic
335 <400> SEQUENCE: 5
336 Met Gly Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr
337 1          5          10          15
339 Leu Thr Val Gln Gln Trp Asp Thr Phe Leu Asn Gly Val Phe Pro Leu
340          20          25          30
342 Asp Arg Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Lys

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/070,099

DATE: 08/12/2002

TIME: 17:41:51

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08122002\J070099.raw

L:10 M:270 C: Current Application Number differs, Replaced Application Number

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date